# High-resolution *in-vivo* human retinal imaging using full-field OCT with optical stabilization of axial motion

**PEDRO MECÊ,**[*] **JULES SCHOLLER, KASSANDRA GROUX, AND CLAUDE BOCCARA**

*Institut Langevin, ESPCI Paris, CNRS, PSL University, 1 rue Jussieu, 75005 Paris, France*
[*]*pedro.mece@espci.fr*

**Abstract:** Time-domain full-field OCT (FF-OCT) represents an imaging modality capable of recording high-speed *en-face* sections of a sample at a given depth. One of the biggest challenges to transfer this technique to image *in-vivo* human retina is the presence of continuous involuntary head and eye axial motion during image acquisition. In this paper, we demonstrate a solution to this problem by implementing an optical stabilization in an FF-OCT system. This was made possible by combining an FF-OCT system, an SD-OCT system, and a high-speed voice-coil translation stage. B-scans generated by the SD-OCT were used to measure the retina axial position and to drive the position of the high-speed voice coil translation stage, where the FF-OCT reference arm is mounted. Closed-loop optical stabilization reduced the RMS error by a factor of 7, significantly increasing the FF-OCT image acquisition efficiency. By these means, we demonstrate the capacity of the FF-OCT to resolve cone mosaic as close as $1.5^o$ from the fovea center with high consistency and without using adaptive optics.

## 1. Introduction

Optical Coherence Tomography (OCT) has become a gold standard for living human retinal imaging in the clinical environment [1,2]. Owing to its unprecedented axial resolution, OCT enables the diagnosis of retinal disorders at the earliest stages and monitoring of the progression of retinal diseases. While physicians are capable of interpreting tomographic retinal cross-sections from OCT, there is nevertheless a demand for *en-face* views. Although OCT can produce *en-face* retinal images from 3-D stacks, it presents a low lateral resolution compared with Adaptive Optics (AO) ophthalmoscopes [3]. By correcting static and dynamic ocular aberrations [4,5], AO can explore the full aperture of the eye's pupil to achieve micrometer lateral resolution. To be able to combine both high axial resolution from OCT, and the high lateral resolution from AO, a great effort was made to develop AO-OCT systems [6,7]. However, the commercialization and the clinical deployment of AO-OCT are still challenging, mainly because of cost, optical complexity, size and heavy image post-processing linked to the use of AO [6].

Another modality of OCT is the time-domain Full-Field OCT (FF-OCT) [8] (henceforth, we refer to this imaging modality simply as FF-OCT). Contrary to conventional OCT, which scans an illumination spot across a tissue surface and records at each spot the depth structure, FF-OCT uses a spatially incoherent light source, high-speed megapixel camera and time-domain phase modulation to acquire *en-face* sections of the sample at a given depth. One attractive point of this technique for retinal imaging, is the fact that FF-OCT presents a low sensitivity to low-order ocular aberrations (*i.e.* defocus and astigmatism), due to the use of the spatially incoherent light source [9,10]. This interesting feature was highlighted in [11], where cone mosaic at $6^o$ from the foveal center of a healthy subject (without prescription eyeglasses) was resolved without using AO. This first image was made possible by the combination of FF-OCT and a Spectral-Domain OCT (SD-OCT). Through the images generated by the SD-OCT both retinal cross-section and

the FF-OCT reference mirror were visible, and a manual match, just before the acquisition, was possible. Nevertheless, this strategy faces significant challenges in providing consistent and reproducible images, mainly due to involuntary axial head and eye motion, which can achieve a median maximum axial speed in the order of 1mm/s [12]. As a consequence, a considerable amount of images has to be acquired to increase the chances of matching the optical path of both arms for a retinal layer of interest.

Hence, to be able to record efficiently and consistently *en-face* sections of the retina *in-vivo*, an axial optical stabilization strategy is necessary. An optical stabilization of axial motion applied in time-domain FF-OCT for retinal imaging was first proposed by Miller *et al.*, combining the FF-OCT with an auxiliary time-domain 1-D OCT and a voice coil translation stage in the reference arm [13]. However, the time-domain 1-D OCT used for axial motion tracking was too slow to work *in-vivo* (20 A-scans per second). Later, Pircher *et al.* [14] combined a transverse scanning OCT system (which directly extracts *en-face* images) with a high-speed Fourier domain 1-D OCT, used to track the cornea apex signal, and a rapid scanning optical delay line montage for real-time optical stabilization. Although the achieved performance was satisfactory, because of the scanning illumination/detection scheme of the proposed OCT, *en-face* images were acquired with a low frame rate, limited FOV and suffered from motion artifacts and distortion due to fixational eye movements [15]. Therefore, continuous involuntary head and eye motion remains one of the biggest challenges in time-domain FF-OCT when imaging the living human retina.

In this paper, we present an FF-OCT retinal imaging system with optical stabilization of axial motion. This system was made possible by combining the FF-OCT system with an SD-OCT. The idea is to use B-scans generated by the SD-OCT to measure the retinal axial position and to drive a high-speed voice-coil translation stage, where the reference arm of the FF-OCT is mounted, correcting for the eye motion. By these means, we demonstrate the capacity of the FF-OCT to resolve cone mosaics as close as $1.5^o$ from the fovea center with high consistency and efficiency, without using adaptive optics.

## 2. Methods

### 2.1. Experimental setup

Figure 1 presents the schematic drawing of the custom-built FF-OCT system coupled with a SD-OCT system. The FF-OCT comprises a light-emitting diode (LED) with $\lambda = 850$ nm center wavelength and 30 nm bandwidth (M850L3, Thorlabs), used as a spatially incoherent illumination source, giving a theoretical axial resolution of approximately 8 $\mu$m in water. The LED is focused by a condenser 20 mm in front of the eye's pupil. A physical diaphragm is positioned in front of the LED, conjugate to the retina and the FF-OCT reference mirror. The illumination beam is split into the reference and the sample arms by a 50:50 cubic beam splitter (BS). For the reference arm, an Olympus 10X/0.25 NA Plan Achromat objective is used with a silicon mirror placed at the focal plane of the objective. The whole reference arm (microscope objective and silicon mirror) is mounted on a fast voice-coil translation stage (X-DMQ12P-DE52, Zaber Technologies Inc.), allowing for adjustment of the coherence gate depth position with a maximum speed of 10 mm/s. For the sample arm, optical windows were introduced to compensate for dispersion effects (not shown in Fig. 1) and the healthy volunteer's eye was aligned along the optical axis. The FF-OCT light beam arrives with an 8 mm diameter in the eye's pupil. The back-scattered photons from both arms are recombined by the same BS and focused onto a high-speed (up to 720 Hz) CMOS camera (Q-2A750-Hm/CXP-6, Adimec) for FF-OCT imaging.

The use of an SD-OCT system coupled with our custom-built FF-OCT is driven by two motivations: firstly, to measure the retinal axial position through SD-OCT cross-sectional images, which is a crucial step to correct for axial motion; secondly, once the axial motion is stabilized, to display in real-time SD-OCT retinal cross-sections of sufficient quality to guide the positioning of the FF-OCT coherence gate at the retinal layer of interest. To address these
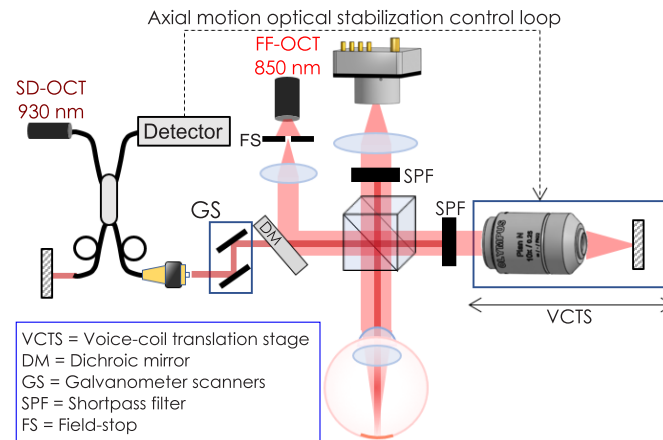
**Fig. 1.** Schematic drawing of the custom-built FF-OCT system coupled with an SD-OCT for real-time axial motion correction and FF-OCT coherence gate positioning guidance.

motivations, we considered a trade-off between acquisition rate, for fast axial motion tracking, and enough signal-to-noise ratio (SNR), for retinal layer identification, helping the positioning of the FF-OCT coherence gate. The SD-OCT system (Thorlabs Ganymede-II) was coupled through a dichroic mirror positioned between the FF-OCT illumination source and the BS. The SD-OCT galvanometer scanners were 20-cm away and not conjugate to the eye's pupil due to three main factors: mechanical restrictions imposed by the commercial SD-OCT system, to avoid specular reflections in the FF-OCT detector by introducing optical lens in the sample arm [16], and to favor a compact optical design.

The SD-OCT system comprises a broadband superluminescent diode with $\lambda = 930$ nm center wavelength and 60 nm bandwidth, giving a theoretical axial resolution of $4.5 \mu m$ in water. It presents an A-scan rate up to 36 kHz with a sensitivity of about 96 dB and 1024 axial pixels, comprising an imaging depth (axial range) of 2.2 mm in water. As the galvanometer scanners and the eye's pupil are not conjugated, the SD-OCT light beam arrives in the eye's pupil with a 4 mm diameter, slightly scanning the pupil plane. Shortpass filters were introduced to block the SD-OCT signal in the reference and detection arms of the FF-OCT.

## 2.2. Axial motion correction

Initially, we manually match the optical path of the FF-OCT reference mirror with a fixed model eye in the sample arm to determine the initial position of the voice-coil translation stage. This calibration step is necessary only once. B-scans generated by the SD-OCT are used to measure the retinal axial position by the following means: 1) A-scans composing the B-scan are averaged, obtaining an axial profile with increased SNR (named from here after "averaged A-scan"); 2) the averaged A-scan is cross-correlated with a reference A-scan and the relative retinal axial position is measured, 3) parabolic fitting is used to obtain a subpixel estimation of the axial retinal position with sub-micrometer precision. The reference A-scan is the first averaged A-scan generated after activating the axial motion correction. This reference is kept until the axial motion correction procedure is deactivated. After the measurement, a corresponding voltage is sent to the voice-coil translation stage, which moves the whole reference arm of the FF-OCT to compensate for the axial motion, re-matching the reference arm with the retinal layer of interest. A custom-developed real-time calculator, comprising core libraries written in MATLAB, controls the axial motion stabilization closed-loop with a 50 Hz loop rate. Control strategy is a PI (proportional-integral) controller with 0.5 gain and the main steps of the chronogram are: SD-OCT B-scan acquisition

(7 ms), SD-OCT B-scan processing (2 ms), axial eye motion measurement (1 ms) and voice-coil translation stage reaction and positioning (10 ms).

## 2.3.　Image acquisition

To obtain statistics on axial motion, we acquired 13 SD-OCT image sequences (7 to 28 s of duration) from three healthy subjects, for a total of 4700 analyzed A-scans. Simultaneous FF-OCT and SD-OCT image acquisition was performed on a young healthy subject (aged 25) presenting a refractive error of $1.5D \times 0.5D \times 160^o$ (spherical × cylindrical × axis of cylindrical errors). Research procedures followed the tenets of the Declaration of Helsinki. Informed consent was obtained from the subject after the nature and possible outcomes of the study were explained. This study was authorized by the appropriate ethics review boards (CPP and ANSM (IDRCB number: 2019-A00942-55)). Subjects were seated in front of the system and stabilized with a chin and forehead rest and asked to fixate a target placed at an infinite focal conjugate. During the imaging session, subjects were wearing their prescription glasses to increase the signal level of both SD-OCT and FF-OCT [10]. Image acquisition was realized in a dark room, maximizing the pupil dilation. Four sets of data were recorded simultaneously during imaging sessions: FF-OCT images, SD-OCT B-scans, retinal axial position and the position of the voice-coil translation stage (equivalent to the position of the FF-OCT reference mirror). Phase modulation was performed by the residual axial motion after optical stabilization (*i.e.* the residual tracking error). FF-OCT images were acquired at 300 Hz using custom-built software [17]. The FF-OCT camera worked close to saturation to use the whole full well capacity, decreasing relative importance of shot noise [18]. Concerning the SD-OCT, we empirically chose to scan $1^o$ FOV of the retina with 256 A-scans, providing a good trade-off between acquisition speed and SNR. The latter had to be sufficient to distinguish most of the retinal layers, crucial for guiding the positioning of the FF-OCT coherence gate to the retinal layer of interest. During image acquisition, the total power entering the eye from the FF-OCT illumination source and the SD-OCT scanning source were respectively 1.3 mW (during 0.25s) and 0.25 mW (continuous scanning), which are below the ocular safety limits established by the ISO standards for group 1 devices.

## 2.4.　Image processing

As detailed in [18], the intensity recorded on a camera pixel is the coherent sum of the reference and sample beams, containing coherent and incoherent terms, expressed as follows:

$$I_{\phi=\phi_N} = \eta \frac{I_0}{4}(R_{eye} + R_{ref} + R_{incoh} + 2\sqrt{R_{eye}R_{ref}}cos(\Delta\phi_N)) \tag{1}$$

where $I_{\phi=\phi_N}$ is the intensity recorded at time point $N$ with a $\phi_N$ phase shift, $\eta$ is the camera quantum efficiency, $I_0$ is the LED's power output (we considered a 50/50 beam-splitter), $R_{eye}$ is the eye reflectivity (*i.e.* the power reflection coefficient), $R_{ref}$ is the reference mirror reflectivity (*i.e.* the power reflection coefficient), $I_{incoh} = R_{incoh}\frac{I_0}{4}$ is the incoherent light detected by the camera, mainly induced by multiple scattering and back-scattering from out-of-focus retinal layers, and $\Delta\phi$ is the phase difference between the reference and the sample back-scattered signals.

From the acquired image sequence, each image was normalized by dividing each image by its mean value. Since the phase was randomly modulated by the residual tracking error, and to eliminate the incoherent terms, we adopted a 2-phase demodulation. The 2-phase demodulation consists of subtracting one image $I_{\phi=\phi_N}$ from the next $I_{\phi=\phi_{N+1}}$ and taking the modulus:

$$I_{2-phase} = |I_{\phi_N} - I_{\phi_{N+1}}| = \eta \frac{I_0}{2}\sqrt{R_{eye}R_{ref}}|cos(\Delta\phi_N) - cos(\Delta\phi_{N+1})| \tag{2}$$

Note that for a 2-phase demodulation strategy it is not possible to un-mix amplitude and phase. Next, images with a very low or absent useful signal, mainly due to an insufficient

phase shift between consecutive images (*i.e.* low value for $|cos(\Delta\phi_N) - cos(\Delta\phi_{N+1})|$), were automatically detected using an intensity-based threshold algorithm, and then excluded from the image sequence. Finally, useful images were registered using a custom-built normalized cross-correlation algorithm, where the image presenting the highest signal level was chosen as the reference.

## 3. Results

### 3.1. Performance of the axial motion stabilization

Figures 2(A) and (B) present two examples of retinal axial position as a function of time. The black line corresponds to the measured axial position of the retina and the red line corresponds to the position of the FF-OCT reference mirror. Fast oscillations and slower drifts of the retinal axial motion are clearly visible. Time points during blink occurrences (black areas) were automatically excluded from the study after applying an intensity-based algorithm to SD-OCT images. Figures 2(C) and (D) illustrate again the axial motion as a function of time, but now applied to the acquired averaged A-scans, where the red rectangle indicates the coherence gate thickness of the FF-OCT system, *i.e.* depth positions where interference takes place. The continuous motion drastically reduces the chances of acquiring an FF-OCT image of a layer of interest (here the inner/outer segment junction - IS/OS). From the collected data set of retinal axial position temporal evolution (excluding blink occurrences), we computed a peak-to-valley amplitude of $244.9 \pm 49.2\mu m$, a mean deviation from zero position of $62.5 \pm 29.5\mu m$ and a root-mean-square (RMS) error of $61.5 \pm 14.2\mu m$. In this condition, the expected efficiency to acquire FF-OCT images of a layer of interest (absolute error inferior to half of the axial resolution, here $4\mu m$), supposing a perfect initial manual match between both arms of the FF-OCT, is $7.4 \pm 1.8$ %.

Figures 2(A) and (B) also show the difference (blue line) between measured retinal position and FF-OCT reference mirror position. Slower drifts were filtered out and fast oscillations were reduced. An absolute residual tracking error of $6.7 \pm 3.1\mu m$ remains mainly because of the inherent time delay of the control loop and the inertia of the voice-coil translation stage. However, the residual error is low enough to considerably increase the chances of acquiring FF-OCT of a retinal layer of interest, as illustrated in Figs. 2(E) and (F). The presented axial motion stabilization method reduced peak-to-valley amplitude to $48 \pm 2\mu m$, mean deviation from zero position to $0.31 \pm 0.13\mu m$, and RMS error to $9.1 \pm 0.68\mu m$. In this condition, the expected efficiency was increased by a factor of 8.5 ($62.4 \pm 4.1$ %). Visualization 1 shows a simulation of the expected B-scan temporal evolution with optical stabilization compared to the position of the FF-OCT coherence gate. Table 1 summarizes the achieved performance with the proposed method to optically stabilize the axial motion.

**Table 1. Axial motion statistics with and without the use of the method to optically stabilize retinal axial motion. Statistics took into account 13 image sequences of 7 to 28s duration from three subjects, for a total of 4700 analyzed A-scans. PV stands for peak-to-valley amplitude.**

|                        | PV ($\mu m$)   | Mean deviation ($\mu m$) | RMS error ($\mu m$) | Efficiency (%) |
| ---------------------- | -------------- | ------------------------ | ------------------- | -------------- |
| Without stabilization  | $244 \pm 49.2$ | $62.5 \pm 29.5$          | $61.5 \pm 14.2$     | $7.4 \pm 1.8$  |
| With stabilization     | $48 \pm 2$     | $0.31 \pm 0.13$          | $9.1 \pm 0.68$      | $62.4 \pm 4.1$ |

To evaluate the temporal performance of the control loop, we computed its rejection transfer function (or sensitivity function) [19], which describes how different axial motion temporal frequencies are attenuated by the optical stabilization control loop. The rejection transfer function is the ratio between temporal power spectral densities (PSD) of the residual tracking error and of the retinal axial position. Figure 3(A) presents an example of the temporal PSD computed for the retinal position (blue line) and residual tracking error (red line). Note that temporal
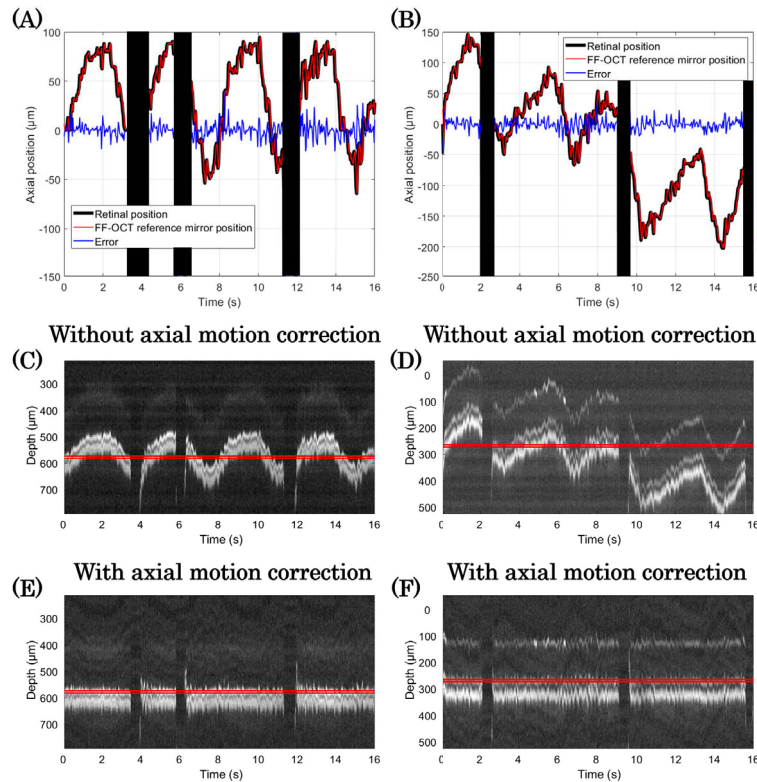
**Fig. 2.** Performance of the axial eye motion optical stabilization. (A) and (B) are two examples of the axial position temporal evolution. Black line: measured axial position of the retina. Red line: position of the FF-OCT reference mirror. Blue line: Residual tracking error. Black areas represent blink occurrences and were excluded from the study. (C) and (D) averaged A-scan time series without axial motion correction. (E) and (F) presents the same averaged A-scan time series as (C) and (D) but now with expected performance after correcting for axial motion. Red rectangles represent the coherence gate axial position and volume. See Visualization 1 for a simulation of the expected B-scan temporal evolution with optical stabilization compared to the position of the FF-OCT coherence gate. Note that all SD-OCT images are shown in linear scale.

frequencies up to approximately 4 Hz are attenuated. Figure 3(B) shows the rejection transfer function obtained by averaging the temporal PSD for different acquired sequences (black line) and the theoretical temporal PSD (red line), assuming a PI controller with 0.5 gain and two-frame delay [19,20]. The good match between the experimental and theoretical plots demonstrates that no additional delay or jitter was present during the acquisition, meaning that the temporal performance of the loop is as expected. Note that, although a loop rate of 50 Hz is adopted, the loop cut-off frequency, *i.e.* the highest corrected temporal frequency (associated with 0 dB), is approximately 4 Hz [3,19].

## 3.2. FF-OCT imaging of living human retina with axial motion correction

Visualization 2 shows an FF-OCT image sequence where the optical stabilization is activated during image acquisition. To begin, the voice-coil translation stage was manually positioned close to the IS/OS junction. A low-signal coming from photoreceptors in the bottom of the frame for a few frames duration is sometimes seen. When optical stabilization is activated, the
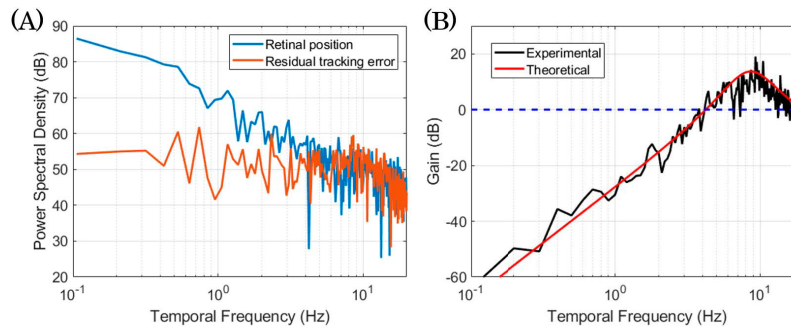
**Fig. 3.** Temporal performance of the axial motion correction control loop. (A) examples of temporal PSD computed for the retinal axial position (blue line) and the residual tracking error (red line). (B) Experimental (black line) rejection transfer function obtained by averaging temporal PSD for different acquired sequences, and the theoretical temporal PSD (red line). Note that, although a loop rate of 50 Hz is adopted, the loop cut-off frequency, *i.e.* the highest corrected temporal frequency (associated with 0 dB), is approximately 4 Hz.

photoreceptor mosaic becomes visible and the signal consistency significantly improves. After activating the axial motion correction, one can observe a temporal fluctuation of the intensity signal where the cone mosaic is visible. This intensity fluctuation occurs mainly due to two phenomena: mispositioning of the coherence gate, or an random phase modulation (see Sect.3.3), both due to the residual tracking error. Figure 4(A) shows the averaged image obtained during axial motion correction. The lateral retinal motion was corrected by digital registration. Another interesting retinal feature perceived in the image is the central foveal depression, *i.e.* an elevation of the IS/OS junction compared to the cone outer segment tip (COST) in the center of the foveola (see schematic of Fig. 4(A) at the right upper part). Because of the depression, this region of the retina is perceived as a dark circled area surrounding the foveal center. This phenomenon happens because this region is out of the FF-OCT coherence gate, which is about $8\mu m$ thick. The diameter of the dark circled area is around $180\mu m$ which corresponds to the typical size of the foveal umbo of a healthy subject [21].

By translating the FF-OCT reference arm laterally (*i.e.* not in depth), perpendicularly to the optical axis, one can generate a tilted coherence gate in the sample arm, producing FF-OCT images with band-type useful FOV. In the case of imaging photoreceptor layers, we expect to obtain two bright bands coming from the IS/OS junction and the COST layer. Tilting the coherence gate allows visualization of the impact of the axial motion correction directly in FF-OCT images. Visualization 3 shows an FF-OCT image sequence with a tilted coherence gate without axial motion correction. It is possible to visualize both photoreceptor bands moving continuously. Band motion is mainly due to axial motion. Lateral retinal motion, caused by fixational eye movements [15], only induces a shift of the retinal zone observed in the band FOV. Visualization 4 shows an FF-OCT image sequence with a tilted coherence gate now with axial motion correction, where band motion is strongly attenuated. In this case, due to the axial motion correction, one can register and average consecutive FF-OCT images to achieve a final image with enhanced quality. Figure 4(B) presents the averaged image generated from Visualization 4. Here, both IS/OS and COST photoreceptor mosaics are visible and could be resolved in a single image.

Figure 4 also shows zoomed areas from (A) and (B) and their respective spatial PSD. The estimated eccentricity (distance from the foveal center) of the zoomed areas are respectively $1^o$, $1.5^o$, $3^o$, $2^o$ and $2^o$. All zoomed areas, except for the first one, presented a clear and well defined Yellot's ring, the spectral signature of photoreceptor mosaic image [22]. Although it seems that
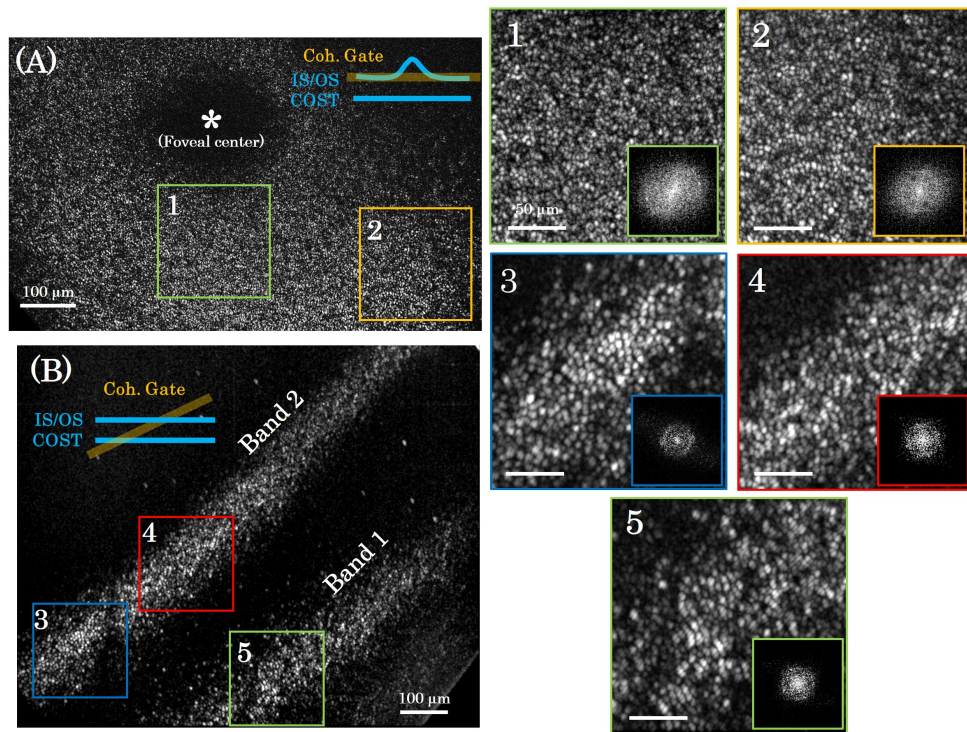
**Fig. 4.** Images of cone photoreceptor mosaic acquired with the help of the axial motion correction method. (A) Cone mosaic averaged image from Visualization 2, acquired at the IS/OS junction at the fovea center. Eccentricities of zoomed areas 1 and 2 are respectively $1^o$, $1.5^o$. (B) Averaged image from Visualization 4. Both cone mosaic from the IS/OS junction and the COST layer close to the fovea center are visible in a single image. Eccentricities of zoomed areas 3, 4 and 5 are respectively $3^o$, $2^o$ and $2^o$. For all zoomed areas the equivalent PSDs were computed.

the subject's pupil diameter was not sufficient to resolve cone mosaic from zoomed area 1, the majority of individual cones are visible. By using modal spacing method proposed by [22], we computed cone densities of 45 000 cones/mm$^2$, 20 700 cones/mm$^2$, 29 000 cones/mm$^2$ and 31 700 cones/mm$^2$ for zoomed areas 2, 3, 4 and 5 respectively. These values are in accordance with the expected cone density for a healthy subject at the eccentricities observed [23].

### 3.3. 2-phase demodulation and FF-OCT image averaging

Typically, a 3- or 4-phase modulation/demodulation strategy makes it possible to eliminate incoherent terms and un-mix the amplitude and the phase information [8]. This strategy works fine for static samples, but it fails for *in-vivo* retinal imaging due to the continuous three-dimensional (3D) eye motion, which randomly modulates the phase. Therefore, we adopted a 2-phase modulation/demodulation strategy, where the residual tracking error directly modulates the signal. Although incoherent terms can still be suppressed, it is not possible to separate amplitude and phase, meaning that the FF-OCT signal presents spatial intensity fluctuations. This random fluctuation, along with shot noise, affects the image contrast (see an example for FF-OCT image - N=1 in Fig. 5(A)).
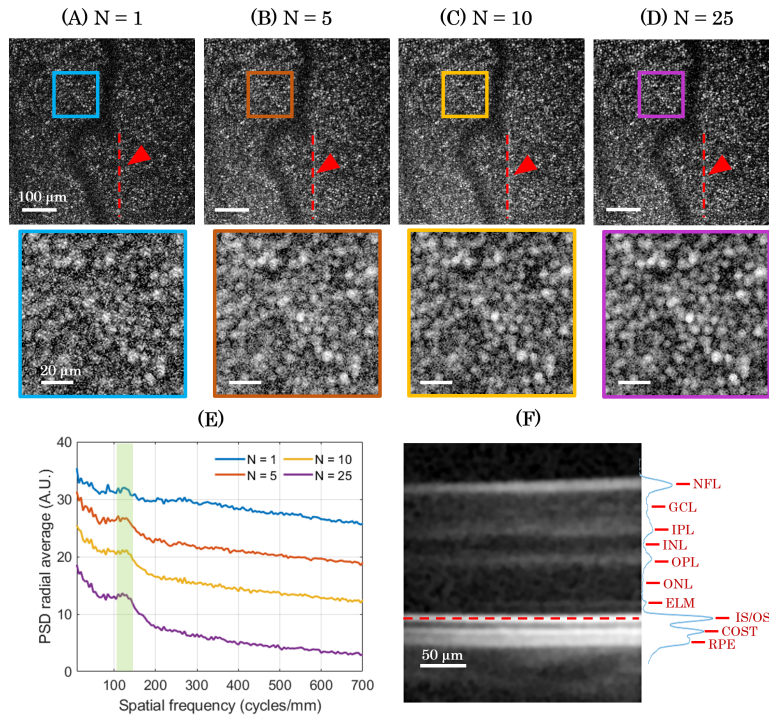
**Fig. 5.** Averaging registered FF-OCT images improves the clarity of photoreceptors. (A),(B),(C) and (D) show images of photoreceptor mosaic at $4^o$ eccentricity for different amount of averaging (from N = 1, N = 5, N = 10 and N = 25 images). The red dashed line, highlighted by a red arrow, indicates where a simultaneous SD-OCT acquisition took place. The increase of image contrast is more visible in the zoomed areas. (E) Power spectral density radial average from each zoomed area. The green area indicates the photoreceptor mosaic spatial frequency, which is present for all the cases presented. (F) Tomographic retinal cross-section image obtained with the SD-OCT after averaging 10 consecutive B-scans. The red dashed line highlights the location where the FF-OCT images were simultaneously acquired, *i.e.* at the IS/OS junction. All averaged images were extracted from Visualization 5.

Since we adopt a random phase modulation, spatial intensity fluctuations vary from one FF-OCT image to another. Therefore, by averaging consecutive FF-OCT images, both phase and noise average out; the signal becomes proportional only to the amplitude, and its contrast is enhanced. This enhancement is highlighted in Fig. 5 for a different number of averaged images (N), where zoomed areas are also shown. Figure 5(E) introduces the computed PSD for each case. Note that for N=1, although the cone mosaic image does not present a sharp contrast, its PSD already indicates a Yellot's ring (green area highlights the cone mosaic spatial frequency), and cone density can be easily computed. The cone density of the zoomed area, localized at $4^o$ eccentricity, was 18 300 cones/mm$^2$, which is in accordance with the expected cone density for a healthy subject at the eccentricity observed [23]. Although image averaging helps to visualize and resolve individual photoreceptors, the appearance of the cone mosaic seems different compared to AO-assisted ophthalmoscopes [3]. This might be explained by the fact that the reflection sites of the IS/OS junction and COST layer might not be exactly within the coherence gate [24].

### 3.4. Simultaneous FF-OCT and SD-OCT imaging of living human retina

Owing to the SD-OCT coupled in our custom-built FF-OCT system, we can access tomographic retinal cross-section images, from the SD-OCT, at the same time as high-resolution *en-face* images from the FF-OCT. Now with the axial motion correction, images from both modalities can be displayed consistently in real-time during the imaging session, helping to position the FF-OCT coherence gate at the retinal layer of interest before image acquisition. Visualization 5 shows one example, where the retinal image sequences from both modalities were recorded simultaneously. We purposely removed the shortpass filter used to block the SD-OCT light source at the FF-OCT camera to help to visualize the SD-OCT acquisition in parallel to the FF-OCT acquisition. The averaged *en-face* image generated by the FF-OCT is shown in Fig. 5(D), where the red dashed line, highlighted by the red arrow, indicates the location where the SD-OCT was acquired simultaneously. The averaged retinal cross-section generated by the SD-OCT is presented in Fig. 5(F). All retinal layers commonly resolved by an SD-OCT were identified and labeled according to [7]. The red dashed line indicates the retinal layer where the FF-OCT image was acquired, *i.e.* the IS/OS junction.

## 4. Discussion

One of the biggest challenges in applying the FF-OCT technique for living human retinal imaging is the presence of continuous involuntary head and eye axial motion during image acquisition. We showed in Sect. 3.1 that axial motion is mainly composed of slow drifts and fast oscillations. Even though *in-vivo* retinal imaging was possible under this condition [11], long and tedious imaging sessions would be necessary to increase the chances of acquiring enough data (the efficiency would be lower than 7%). Moreover, we saw in Sect. 3.3 that, even if photoreceptors could be visualized and resolved in a single frame (see Fig. 5(A)), image averaging is an important process to eliminate the spatial intensity fluctuation imposed by the random phase modulation of FF-OCT signal.

### 4.1. Using SD-OCT to guide FF-OCT in-vivo imaging

In this paper, we present a method to overcome this major limitation, based on the combination of an FF-OCT, an SD-OCT and a high-speed voice-coil translation stage. The SD-OCT is used to measure the retinal axial position and to drive the position of the voice-coil translation stage, where the FF-OCT reference arm is mounted. We showed that the proposed method was able to suppress slow drifts and strongly attenuate fast oscillations, increasing the FF-OCT image acquisition efficiency to almost a factor of 10, *i.e* 60% success. This efficiency was highlighted by Visualization 2, Visualization 4 and Visualization 5, where FF-OCT image sequences with axial motion correction were presented.

    Another essential role of the SD-OCT is to guide the FF-OCT coherence gate positioning. Indeed, once the axial motion is stabilized, and since the initial position of the FF-OCT coherence gate in the sample arm is known (calibration step described in Sect. 2.2), one can manually position the FF-OCT coherence gate assisted by the SD-OCT retinal cross-sections displayed in real-time. In spite of choices of optical design to favor the FF-OCT system, sometimes to the detriment of the SD-OCT system, and the mechanical restrictions imposed by the SD-OCT commercial system, we showed in Fig. 5 and Visualization 5 that retinal cross-section images, with sufficient contrast to identify all retinal layers, can be generated simultaneously to FF-OCT, a crucial condition to guide FF-OCT image acquisition.

### 4.2. High cellular resolution in-vivo retinal imaging with FF-OCT

Throughout this paper, we demonstrated the ability of our FF-OCT imaging system to achieve high cellular resolution in the living human retina. There has been increased interest in the

past ten years in associating different retinal imaging modalities to achieve a cellular resolution in all three dimensions, with most groups combining AO-OCT and AO-SLO [25,26]. The main difference of our multimodal imaging system compared to previous works is the potential to achieve cellular resolution in all three dimensions without AO, significantly reducing the system complexity, size and cost. Figure 4 demonstrated the capacity of the FF-OCT to resolve *in-vivo* cone photoreceptors as close as $1.5^o$ from the foveal center without using any optical aberration compensation other than prescribed eyeglasses. Moreover, the presented FF-OCT has the capacity to generate *en-face* retinal images with an enhanced axial resolution compared to confocal AO-SLO, $8\mu m$ for the former against $40\mu m$ for the latter [13]. Some studies have demonstrated that, for healthy subjects with good eye optics, individual cone photoreceptors could be resolved within SD-OCT or confocal SLO images without optically correcting for ocular aberrations [14,27]. However, most of these studies used illumination/detection scanning techniques, which present an inherent low frame rate, limited FOV and high sensitivity to eye motion, the latter leading to intraframe distortion. All these drawbacks are minimized when using a full-field technique [3,28].

### 4.3.  Limitations

The presented FF-OCT/SD-OCT imaging system still faces some limitations. Firstly, to favor the FF-OCT system, the capacity of the SD-OCT imaging was diminished by not-conjugating the galvanometer scanners to the eye's pupil, affecting its signal level and increasing its sensitivity to the subject alignment. Secondly, even if FF-OCT presents a low sensitivity to low-order ocular aberrations in terms of resolution, the latter still affects the FF-OCT signal level [10]. Ocular aberrations and the variation of the subject pupil diameter might explain the fact that single-frame images obtained in Visualization 2 (when axial motion correction is activated) present a higher signal than Fig. 5(A). Finally, because ocular aberrations affect the SNR of the current FF-OCT, imaging other retinal layers, such as ganglion cells, possible with AO-OCT systems [29], becomes challenging. As discussed in [10], now that FF-OCT images can be acquired with high consistency, a compact implementation of AO could be used, consisting of a wavefront sensorless approach, and a multi-actuator adaptive lens [30] positioned just in front of the eye, like regular glasses (*i.e.* without strict conjugation). As ocular aberrations are dominated by low-order aberrations [4], we expect to significantly improve the FF-OCT SNR, and therefore hope to be able to image less reflective retinal layers.

### 4.4.  Towards dynamic FF-OCT in the living human eye

One interesting way to obtain a high image contrast using FF-OCT is by exploring temporal fluctuations of the backscattered light, revealing subcellular structures and transparent tissues [18,31]. So far, this technique, when applied to retinal imaging, has only been shown for *ex-vivo* retinal imaging. To be able to achieve dynamic *in-vivo* retinal imaging an axial motion correction better than the axial resolution of the FF-OCT system is necessary [32]. Our proposed method achieves this performance 60% of the time, and it might be sufficient. Otherwise, a higher performance would only be possible with a faster reaction of the voice-coil translation stage or with an SD-OCT with a higher acquisition rate (here we are limited to 36 000 A-scans per second).

## 5.  Conclusion

One of the biggest challenges of time-domain FF-OCT when applied for *in-vivo* human retinal imaging is the presence of continuous involuntary head and eye axial motion. We presented a method to optically compensate for axial motion in real-time for time-domain FF-OCT. This was possible by introducing an SD-OCT and a voice-coil translation stage in the system, the former to generate retinal cross-section images, where retinal axial motion could be measured,

and the latter to correct the FF-OCT reference arm position. The closed-loop optical stabilization reduced the RMS error by a factor of 7, significantly increasing the FF-OCT image acquisition efficiency. By these means, we demonstrated the capacity of the FF-OCT to resolve cone mosaic as close as $1.5^o$ from the fovea center with high consistency and repeatability, without the need for AO. The good precision and efficiency achieved in the FF-OCT system, by virtue of the real-time axial motion optical stabilization, together with its relative simplicity and low-cost compared to the other high-resolution techniques which use conventional AO, may pave the way towards the adoption of FF-OCT as a routine clinical imaging system.

## Funding

## Acknowledgments

## Disclosures

The authors declare no conflicts of interest.

## References

1. D. Huang, E. Swanson, C. Lin, J. Schuman, W. Stinson, W. Chang, M. Hee, T. Flotte, K. Gregory, C. Puliafito, and J. G. Fujimoto, "Optical coherence tomography," Science **254**(5035), 1178–1181 (1991).
2. W. Drexler and J. G. Fujimoto eds., *Optical Coherence Tomography: Technology and Applications* (Springer, 2015), 2nd ed.
3. E. Gofas-Salas, P. Mecê, C. Petit, J. Jarosz, L. M. Mugnier, A. M. Bonnefois, K. Grieve, J. Sahel, M. Paques, and S. Meimon, "High loop rate adaptive optics flood illumination ophthalmoscope with structured illumination capability," Appl. Opt. **57**(20), 5635–5642 (2018).
4. J. Jarosz, P. Mecê, J.-M. Conan, C. Petit, M. Paques, and S. Meimon, "High temporal resolution aberrometry in a 50-eye population and implications for adaptive optics error budget," Biomed. Opt. Express **8**(4), 2088–2105 (2017).
5. P. Mecê, E. Gofas-Salas, C. Petit, F. Cassaing, J. Sahel, M. Paques, K. Grieve, and S. Meimon, "Higher adaptive optics loop rate enhances axial resolution in nonconfocal ophthalmoscopes," Opt. Lett. **44**(9), 2208–2211 (2019).
6. R. S. Jonnal, O. P. Kocaoglu, R. J. Zawadzki, Z. Liu, D. T. Miller, and J. S. Werner, "A review of adaptive optics optical coherence tomography: technical advances, scientific applications, and the future," Invest. Ophthalmol. Visual Sci. **57**(9), OCT51–OCT68 (2016).
7. M. Pircher and R. J. Zawadzki, "Review of adaptive optics oct (AO-OCT): principles and applications for retinal imaging," Biomed. Opt. Express **8**(5), 2536–2562 (2017).
8. A. Dubois, K. Grieve, G. Moneron, R. Lecaque, L. Vabre, and C. Boccara, "Ultrahigh-resolution full-field optical coherence tomography," Appl. Opt. **43**(14), 2874–2883 (2004).
9. O. Thouvenin, K. Grieve, P. Xiao, C. Apelian, and A. C. Boccara, "En face coherence microscopy," Biomed. Opt. Express **8**(2), 622–639 (2017).
10. P. Mecê, P. Xiao, V. Mazlin, J. Scholler, K. Grieve, J.-A. Sahel, M. Fink, and C. Boccara, "Towards lens-based wavefront sensorless adaptive optics full-field oct for in-vivo retinal imaging (conference presentation)," in *Optical Coherence Tomography and Coherence Domain Optical Methods in Biomedicine XXIII*, vol. 10867 (International Society for Optics and Photonics, 2019), p. 1086722.
11. P. Xiao, V. Mazlin, K. Grieve, J.-A. Sahel, M. Fink, and A. C. Boccara, "In vivo high-resolution human retinal imaging with wavefront-correctionless full-field oct," Optica **5**(4), 409–412 (2018).
12. L. Ginner, A. Kumar, D. Fechtig, L. M. Wurster, M. Salas, M. Pircher, and R. A. Leitgeb, "Noniterative digital aberration correction for cellular resolution retinal optical coherence tomography in vivo," Optica **4**(8), 924–931 (2017).
13. D. T. Miller, J. Qu, R. S. Jonnal, and K. E. Thorn, "Coherence gating and adaptive optics in the eye," in *Coherence Domain Optical Methods and Optical Coherence Tomography in Biomedicine VII*, vol. 4956 (International Society for Optics and Photonics, 2003), pp. 65–72.
14. M. Pircher, E. Götzinger, H. Sattmann, R. A. Leitgeb, and C. K. Hitzenberger, "In vivo investigation of human cone photoreceptors with slo/oct in combination with 3d motion correction on a cellular level," Opt. Express **18**(13), 13935–13944 (2010).
15. P. Mecê, J. Jarosz, J.-M. Conan, C. Petit, K. Grieve, M. Paques, and S. Meimon, "Fixational eye movement: a negligible source of dynamic aberration," Biomed. Opt. Express **9**(2), 717–727 (2018).

16. E. Auksorius and A. C. Boccara, "Dark-field full-field optical coherence tomography," Opt. Lett. **40**(14), 3272–3275 (2015).
17. J. Scholler, "FFOCT control and acquisition software," (2019). https://doi.org/10.5281/zenodo.3137245.
18. J. Scholler, V. Mazlin, O. Thouvenin, K. Groux, P. Xiao, J.-A. Sahel, M. Fink, C. Boccara, and K. Grieve, "Probing dynamic processes in the eye at multiple spatial and temporal scales with multimodal full field oct," Biomed. Opt. Express **10**(2), 731–746 (2019).
19. F. Roddier, *Adaptive Optics in Astronomy* (Cambridge University Press, 1999).
20. S. Meimon, C. Petit, T. Fusco, and C. Kulcsar, "Tip–tilt disturbance model identification for kalman-based control scheme: application to xao and elt systems," J. Opt. Soc. Am. A **27**(11), A122–A132 (2010).
21. S. Duke-Elder, "The anatomy of the visual system," A System of Ophthalmology **2**, 363–382 (1961).
22. R. F. Cooper, C. S. Langlo, A. Dubra, and J. Carroll, "Automatic detection of modal spacing (yellott's ring) in adaptive optics scanning light ophthalmoscope images," Ophthalmic Physiol. Opt. **33**(4), 540–549 (2013).
23. L. Sawides, A. de Castro, and S. A. Burns, "The organization of the cone photoreceptor mosaic measured in the living human retina," Vision Res. **132**, 34–44 (2017).
24. M. Pircher, B. Baumann, E. Götzinger, H. Sattmann, and C. K. Hitzenberger, "Simultaneous slo/oct imaging of the human retina with axial eye motion correction," Opt. Express **15**(25), 16922–16932 (2007).
25. R. J. Zawadzki, S. M. Jones, S. Pilli, S. Balderas-Mata, D. Y. Kim, S. S. Olivier, and J. S. Werner, "Integrated adaptive optics optical coherence tomography and adaptive optics scanning laser ophthalmoscope system for simultaneous cellular resolution in vivo retinal imaging," Biomed. Opt. Express **2**(6), 1674–1686 (2011).
26. Z. Liu, J. Tam, O. Saeedi, and D. X. Hammer, "Trans-retinal cellular imaging with multimodal adaptive optics," Biomed. Opt. Express **9**(9), 4246–4262 (2018).
27. N. D. Shemonski, F. A. South, Y.-Z. Liu, S. G. Adie, P. S. Carney, and S. A. Boppart, "Computational high-resolution optical imaging of the living human retina," Nat. Photonics **9**(7), 440–443 (2015).
28. P. Bedggood and A. Metha, "De-warping of images and improved eye tracking for the scanning laser ophthalmoscope," PLoS One **12**(4), e0174617 (2017).
29. Z. Liu, K. Kurokawa, F. Zhang, J. J. Lee, and D. T. Miller, "Imaging and quantifying ganglion cells and other transparent neurons in the living human retina," Proc. Natl. Acad. Sci. **114**(48), 12803–12808 (2017).
30. S. Bonora, Y. Jian, P. Zhang, A. Zam, E. N. Pugh, R. J. Zawadzki, and M. V. Sarunic, "Wavefront correction and high-resolution in vivo oct imaging with an objective integrated multi-actuator adaptive lens," Opt. Express **23**(17), 21931–21941 (2015).
31. C. Apelian, F. Harms, O. Thouvenin, and A. C. Boccara, "Dynamic full field optical coherence tomography: subcellular metabolic contrast revealed in tissues by interferometric signals temporal analysis," Biomed. Opt. Express **7**(4), 1511–1524 (2016).
32. J. Scholler, "Motion artifact removal and signal enhancement to achieve in vivo dynamic full field OCT," Opt. Express **27**(14), 19562–19572 (2019).